# TRANSMITTAL FORM

*(to be used for all correspondence after initial filing)*

| | |
|---|---|
| **Application Number** | 09/287,296 |
| **Filing Date** | April 7, 1999 |
| **First Named Inventor** | Yeogirl YUN et al. |
| **Group Art Unit** | 3624 |
| **Examiner Name** | T. T. Havan |

| Total Number of Pages in This Submission | | **Attorney Docket Number** | 002566-013000 |
|---|---|---|---|

## ENCLOSURES *(check all that apply)*

☐ Fee Transmittal Form

　☐ Fee Attached

☐ Amendment / Reply

　☐ After Final

　☐ Affidavits/declaration(s)

☐ Extension of Time Request

☐ Express Abandonment Request

☐ Information Disclosure Statement

☐ Certified Copy of Priority Document(s)

☐ Response to Missing Parts/ Incomplete Application

　☐ Response to Missing Parts under 37 CFR 1.52 or 1.53

☐ Assignment Papers *(for an Application)*

☐ Drawing(s)

☐ Declaration and Power of Attorney

☐ Licensing-related Papers

☐ Petition

☐ Petition to Convert to a Provisional Application

☐ Power of Attorney, Revocation Change of Correspondence Address

☐ Terminal Disclaimer

☐ Request for Refund

☐ CD, Number of CD(s) _____

☐ After Allowance Communication to Group

☐ Appeal Communication to Board of Appeals and Interferences

☒ Appeal Communication to Group (Appeal Notice, Brief, Reply Brief)
**Amended Appeal Brief**

☐ Proprietary Information

☐ Status Letter

☐ Application Data Sheet

☐ Request for Corrected Filing Receipt with Enclosures

☐ A self-addressed prepaid postcard for acknowledging receipt

☐ Other Enclosure(s) *(please identify below)*:

| Remarks | ☒ The Commissioner is hereby authorized to charge any additional fees required or credit any overpayments to Deposit Account No. 19-2380 for the above identified docket number. |
|---|---|

## SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

| Firm *or* Individual name | Daniel S. Song; Reg. No. 43,143 Nixon Peabody LLP 401 9th Street, N.W. Suite 900 Washington, D.C. 20004-2128 |
|---|---|
| Signature | |
| Date | March 28, 2007 |

## CERTIFICATE OF MAILING OR TRANSMISSION [37 CFR 1.8(a)]

I hereby certify that this correspondence is being:

　☐ deposited with the United States Postal Service on the date shown below with sufficient postage as first class mail in an envelope addressed to: Mail Stop _____, Commissioner for Patents, P. O. Box 1450, Alexandria, VA 22313-1450

　☐ transmitted by facsimile on the date shown below to the United States Patent and Trademark Office at (571) _____.

| Date | Signature |
|---|---|
| | Typed or printed name |

10201072.2

# IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
## BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

| | |
|---|---|
| In re PATENT application of | ) |
| Yeogirl YUN et al. | ) Confirmation No. 3718 |
| Serial No. 09/287,296 | ) Group Art Unit: 3624 |
| Filed: April 7, 1999 | ) Examiner: T. T. Havan |
| For: METHOD AND APPARATUS FOR DEFINING DATA OF INTEREST | ) |
| | ) Date: March 28, 2007 |

## AMENDED APPEAL BRIEF

Mail Stop Appeal Brief- Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

This Amended Appeal Brief is submitted in support of the Notice of Appeal filed September 22, 2006, and further addresses the deficiencies noted in the Notification of Non-Compliant Appeal Brief mailed March 21, 2007. In particular, Evidence Appendix and Related Proceedings Appendix are now submitted on new pages. Acceptance of the Amended Appeal Brief is respectfully requested.

### I. REAL PARTY IN INTEREST

CNET Networks, Inc. is the real party in interest.

### II. RELATED APPEALS AND INTERFERENCES

There are presently no appeals or interferences known to the Appellants, the Appellants' representative, or the assignee, which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

### III. STATUS OF THE CLAIMS

Claims 1, 2, 4-7, 9-21 and 23-34 are pending in the present application, as submitted in an Amendment filed January 4, 2006, in response to the Office Action mailed September 21, 2005. This Appeal is taken from the rejection of claims 1, 2, 4-

7, 9-21 and 23-34, the claims being submitted in the CLAIMS APPENDIX submitted herewith.

## IV.    STATUS OF AMENDMENTS

No amendments have been filed subsequent to the Final Office Action mailed March 24, 2006. A Request for Reconsideration without any claim amendments was filed on June 22, 2006 in response to the Final Office Action, which was considered by the Examiner, but the rejection maintained, in the Advisory Action mailed August 16, 2006.

## V.    SUMMARY OF CLAIMED SUBJECT MATTER

The present invention relates to a method and apparatus for facilitating extraction of data of interest from a plurality of web sites. Conventionally, "web crawlers" are created by computer programmers to retrieve information from a particular web site, for example, to extract desired information for a category of products from on-line merchants for use in an electronic catalog. However, different web sites, for example, web sites of different on-line merchants, utilize different data structures. There is no standardized structure, method or protocol for presenting and storing information or data among different web sites that is uniformly followed by different on-line merchants. In addition, each web site generally utilizes a plurality of web pages in the web site to which a user has to navigate to obtain the desired data of interest regarding a product available through the web site, for example.

A particular crawler that is created by a computer programmer for use in extraction of data from one web site generally cannot be used to extract data from other web sites due to the variations in data structure, method and/or protocol implemented by other web sites. Thus, a new crawler must be created by a computer programmer to extract data for each web site, the creation of new crawlers being time consuming and expensive. Consequently, extracting data of interest, for example, regarding a particular product from a plurality of different web sites such as merchant web sites, can be extremely difficult, expensive, and time consuming.

The present invention provides a novel method and system for extracting data of interest from a plurality of web sites that greatly facilitates the extraction process

by providing tools that can be used, even by non-programmers, to extract desired information from the plurality of web sites. (See Pg. 3, lines 9-10; Pg. 8, lines 19-23). More specifically, the present invention allows the user to generate extraction patterns directly from the output of the web site itself, such as the HTML source view of a web browser, so that other desired information can also be extracted from the web site using the generated extraction patterns. (See Pg. 3, lines 13-20; Figs. 10, 14-18 and related disclosure in Pg. 23, lines 13-19; Pg. 24, line 18-Pg. 26, line 6). As explained to the Examiner on numerous instances, extraction patterns of the present invention and recited in the present claims are not keyword queries. Thus, the present invention can be used to generate extraction patterns easily based on the output of the web site itself. (See Figures 14 to 17; Page 24, line 18 to Page 25, line 24, that describe example implementation of the present invention and development of extraction patterns). A value can then be used in conjunction with the developed extraction pattern to extract different data of interest from the particular web site. Correspondingly, the present invention allows facilitated extraction of desired data of interest from a plurality of web sites in a rapid, cost effective manner, without requiring a programmer to create a crawler for each web site from which data of interest is desired.

Accordingly, independent claim 1 recites a method of extracting data of interest from a plurality of web sites, the method comprising, for each respective web site W in said plurality of web sites, creating a respective description of data of interest that identifies the web site W (Pg. 5, lines 1-3; Fig. 3, 300, 306; Pg. 12, lines 13-15), developing an extraction pattern from a web page output from the respective web site using a graphical user interface tool (Pg. 5, lines 3-5; Fig. 1, 114; Pg. 7, lines 9-14; Pg. 8, lines 19- 23; Fig. 4, 404; Pg. 14, lines 14-15; Pg. 15, lines 8-12; Pg. 16, lines 7-8; Fig. 10, 1002; Pg. 23, lines 13-17; Fig. 15, 1502, Pg. 25, lines 1-4; Fig. 16, 1602; Pg. 25, lines 5-7), the extraction pattern being adapted to identify at least a portion of an output of a web site and to extract information from a plurality of web pages of the respective web site W (Fig. 15; Pg. 24, line 15-Pg. 25, line 4; Fig. 23; Pg. 27, lines 14-16), and associating the developed extraction pattern with the respective description of data of interest for the respective web site W (Fig. 4, 408; Pg. 18, lines

5-8; Fig. 12; Pg. 24, lines 11-14; Fig. 12; Pg. 25, lines 9-24; Pg. 26, lines 18-26; Figs. 20 and 21). Independent claim 1 further recites receiving a value for use as an extraction parameter for the developed extraction patterns (Fig. 2, 204; Pg. 10, line 23-Pg. 11, line 7; Pg 24, lines 19-25), and obtaining the data of interest by querying web sites in the plurality of web sites using the value and the extraction patterns associated with the respective descriptions of data of interest (Pg. 9, lines 10-12; Fig. 2, 208; Pg. 11, lines 12-17; Fig. 4, 408; Pg. 18, lines 2-5), wherein when the data of interest includes data of interest from at least two web sites of the plurality of web sites, the data of interest from the at least two web sites is extracted (Pg. 4, lines 13-14; Pg. 11, lines 18-24).

Claim 2 is dependent on claim 1 and specifically recites that the graphical user interface tool includes a web browser (Pg. 14, lines 13-15; Fig. 9; Pg. 22, line 22-Pg. 23, line 2).

Claim 4 is dependent on claim 2 and specifically recites applying the extraction pattern to the output of the web site that is displayed in a source view in the web browser thereby identifying the at least a portion of the output for the web site (Pg. 16, lines 7-8; Fig. 10, 1002; Pg. 23, lines 13-19), and displaying a rendered version of the at least a portion of the output of the web site (Pg. 16, lines 8-9; Fig. 14; Pg. 24, lines 18-22).

Claim 5 is also dependent on claim 2 and specifically recites that the graphical user interface tool further includes a plurality of predefined extraction patterns (Fig. 10, 908; Pg. 23, lines 7-14).

Claim 6 which is dependent on claim 5 further recites that the plurality of predefined extraction patterns includes at least one of an extraction pattern for matching a hyperlink, an extraction pattern for matching a form, and an extraction pattern for matching a price (Fig. 10, 908).

Claim 9 that is dependent on claim 1 recites that the developing of an extraction pattern includes receiving a selection of an extraction command from a predetermined list of extraction commands (Fig. 4, 406; Pg. 16, lines 13-16; Pg. 17, Table 1).

Claim 12 is dependent on claim 9 and recites applying a test condition comprising a logical test for at least one corresponding argument, and that the respective description of data of interest continues executing when the logical test is satisfied (Fig.4; 414; Pg. 19, lines 3-6).

Claim 14 is dependent on claim 12 and recites that the test condition further comprises a result code that returns an error when the output of the respective web site has changed (Pg. 20, lines 20-25).

Claim 15 is also dependent on claim 12 and recites that the test condition further comprises a result code that returns an error when the output of the respective web site has no information about the product (Pg. 20, lines 14-19).

Claim 16 is dependent on claim 9 and recites that the predetermined list of extraction commands includes an extraction command for segmenting the output of the respective web site into a plurality of units, each of the plurality of units matching the extraction pattern (Pg. 19, lines 20-24).

Independent claim 18 recites an apparatus for extracting information of interest from a plurality of web sites (Fig. 1, 104, 106, 112, 108; Pg. 7, lines 6-8), the apparatus comprising, for each respective web site W in the plurality of web sites, a means for creating a respective description of data of interest that identifies the web site W (Pg. 5, lines 1-3; Fig. 1, 110; Pg. 7, lines 8-9; Fig. 3, 300, 306; Pg. 12, lines 13-15), and a means for developing an extraction pattern from a web page output from the respective web site using a graphical user interface tool (Pg. 5, lines 3-5; Fig. 1, 114; Pg. 7, lines 9-14; Pg. 8, lines 19- 23; Fig. 4, 404; Pg. 14, lines 14-15; Pg. 15, lines 8-12; Pg. 16, lines 7-8; Fig. 10, 1002; Pg. 23, lines 13-17; Fig. 15, 1502, Pg. 25, lines 1-4; Fig. 16, 1602; Pg. 25, lines 5-7), the extraction pattern being adapted to extract data from a plurality of web pages of the web site (Fig. 15; Pg. 24, line 15-Pg. 25, line 4; Fig. 23; Pg. 27, lines 14-16), and a means for associating the developed extraction pattern with the respective description of data of interest for the respective web site W (Fig. 1, 114; Fig. 4, 408; Pg. 18, lines 5-8; Fig. 12; Pg. 24, lines 11-14; Fig. 12; Pg. 25, lines 9-24; Pg. 26, lines 18-26; Figs. 20 and 21). Independent claim 18 further recites a means for receiving a value for use as an extraction parameter in the developed extraction patterns (Fig. 1, 100, 108; Fig. 2, 204; Pg. 10, line 23-Pg. 11,

line 7; Pg 24, lines 19-25), and a means for obtaining the data of interest by querying web sites in the plurality of web sites using the value and the developed extraction patterns associated with the respective descriptions of data of interest (Fig. 1, 108; Pg. 9, lines 10-12; Fig. 2, 208; Pg. 11, lines 12-17; Fig. 4, 408; Pg. 18, lines 2-5), wherein, when the data of interest includes data from at least two web sites of the plurality of web sites, the means for obtaining the data of interest provides the data of interest from the at least two web sites (Pg. 4, lines 13-14; Pg. 11, lines 18-24).

Claim 19 is dependent on claim 18 and recites that the means for developing an extraction pattern includes means for selecting an instruction from a predetermined list of instructions (Fig. 4, 406; Pg. 16, lines 13-16; Pg. 17, Table 1).

Claim 20 is also dependent on claim 18 and recites that the graphical user interface tool includes a web browser (Pg. 14, lines 13-15; Fig. 9; Pg. 22, line 22-Pg. 23, line 2).

Independent claim 21 recites a computer data signal embodied in a carrier wave including a software module for creating a description of data of interest (Pg. 5, lines 1-3; Fig. 1, 108, 110; Pg. 7, lines 7-9; Fig. 3, 300, 306; Pg. 12, lines 13-15), the software module including, a set of operations for interactively developing an extraction pattern from a web page output of a target web site using a graphical user interface tool (Pg. 5, lines 3-5; Fig. 1, 114; Pg. 7, lines 9-14; Pg. 8, lines 19- 23; Fig. 4, 404; Pg. 14, lines 14-15; Pg. 15, lines 8-12; Pg. 16, lines 7-8; Fig. 10, 1002; Pg. 23, lines 13-17; Fig. 15, 1502, Pg. 25, lines 1-4; Fig. 16, 1602; Pg. 25, lines 5-7), the developed extraction pattern being adapted to extract data of interest from a plurality of web pages of the target web site (Fig. 15; Pg. 24, line 15-Pg. 25, line 4; Fig. 23; Pg. 27, lines 14-16). Claim 21 also recites that the software module also includes a set of operations for receiving a selection of an instruction from a predefined set of instructions for inclusion in the description of data of interest (Fig. 4, 406; Pg. 16, lines 13-16; Pg. 17, Table 1), a set of operations for associating the extraction pattern with the instruction (Fig. 4, 408; Pg. 18, lines 5-8; Fig. 12; Pg. 24, lines 11-14; Fig. 12; Pg. 25, lines 9-24; Pg. 26, lines 18-26; Figs. 20 and 21), and a set of operations for testing the instruction using the extraction pattern and the contents of a buffer (Fig.4; 414; Pg. 19, lines 3-6), wherein the buffer includes a portion of the web page output

of the web site associated with the description of data of interest (Pg. 16, lines 7-8; Fig. 10, 1002; Pg. 23, lines 13-19). Independent claim 21 further recites a software module for using the description of data of interest to obtain data of interest from the target web site when a value for use as an extraction parameter for the developed extraction pattern is provided (Fig. 2, 204, 208; Pg. 9, lines 10-12; Pg. 10, line 23-Pg. 11, line 17; Pg 24, lines 19-25; Fig. 4, 408; Pg. 18, lines 2-5).

Claim 23 which is dependent on independent claim 1 recites that the extraction pattern comprises a pre-condition regular expression, a portion of data of interest regular expression, and a post-condition regular expression and that the operations for interactively developing an extraction pattern comprises refining at least one of the pre-condition regular expression, the portion of data of interest regular expression, and the post-condition regular expression (Pg. 15, lines 13-24).

Claim 27 is dependent on independent claim 18 and further recites that the extraction pattern comprises a pre-condition regular expression, a portion of data of interest regular expression, and a post-condition regular expression and that the means for developing comprise refining at least one of the pre-condition regular expression, the portion of data of interest regular expression, and the post-condition regular expression (Pg. 15, lines 13-24).

Claim 28 is dependent on independent claim 21 and further recites that the extraction pattern comprises a pre-condition regular expression, a portion of data of interest regular expression, and a post-condition regular expression and that the operations for developing comprise refining at least one of the pre-condition regular expression, the portion of data of interest regular expression, and the post-condition regular expression (Pg. 15, lines 13-24).

Independent claim 32 recites a computer implemented method of obtaining data of interest from a plurality of web sites comprising, developing a description of data of interest for each web site in said plurality of web sites from web page output from the plurality of web sites using a graphical user interface tool that includes a web browser, each respective description of data of interest specifying an address for a corresponding web site in the plurality of web sites (Pg. 5, lines 1-5; Fig. 1, 114; Pg. 7, lines 9-14; Pg. 8, lines 19-23; Fig. 3, 300, 306; Pg. 12, lines 13-15; Fig. 4, 404; Pg.

14, lines 13-15; Pg. 15, lines 8-12; Pg. 16, lines 7-8; Fig. 9; Pg. 22, line 22-Pg. 23, line 2; Fig. 10, 1002; Pg. 23, lines 13-17; Fig. 15, 1502, Pg. 25, lines 1-4; Fig. 16, 1602; Pg. 25, lines 5-7). Claim 32 also recites that each respective description of data of interest includes an extraction pattern adapted to identify at least a portion of the output of a web site and to extract user specified information from a plurality of web pages of the corresponding web site (Fig. 2, 204, 208; Pg. 9, lines 10-12; Pg. 10, line 23-Pg. 11, line 17; Fig. 4, 408; Pg. 18, lines 2-5; Fig. 15; Pg. 24, line 15-Pg. 25, line 4; Fig. 23; Pg. 27, lines 14-16). Independent claim 32 further recites receiving a value for use as an extraction parameter for the developed extraction patterns (Fig. 2, 204; Pg. 10, line 23-Pg. 11, line 7; Pg 24, lines 19-25), and obtaining said data of interest by querying web sites in the plurality of web sites using the value and the extraction patterns in the respective descriptions of data of interest (Pg. 9, lines 10-12; Fig. 2, 208; Pg. 11, lines 12-17; Fig. 4, 408; Pg. 18, lines 2-5).

Claim 33 dependent on claim 32 further recites that each extraction pattern comprises a pre-condition regular expression, a portion of data of interest regular expression, and a post-condition regular expression and that the step of developing a description of data of interest comprises refining at least one of the pre-condition regular expression, the portion of data of interest regular expression, and the post-condition regular expression (Pg. 15, lines 13-24).

## VI.  GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

In the Final Office Action mailed March 24, 2006, the Examiner has maintained the sole ground for rejection of claims 1, 2, 4-7, 9-21 and 23-34 under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent No. 6,094,649 to Bowen et al. in view of U.S. Patent No. 6,144,991 to England. Thus, the issue on appeal is whether claims 1, 2, 4-7, 9-21 and 23-34 are unpatentable under 35 U.S.C. 103(a).

## VII.  ARGUMENT

As set forth herein below, the Appellants respectfully contend that the Examiner has failed to establish a *prima facie* case of obviousness, and the rejection set forth in the final Office Action mailed March 24, 2006 should be reversed.

It is initially noted that the final Office Action of March 24, 2006 merely summarily reaffirmed the Examiner's prior Office Action of September 21, 2005 in which all of the pending claims were substantively rejected under 35 U.S.C. 103(a). Correspondingly, Examiner's summary rejection set forth in the final Office Action is briefly addressed herein below and the Examiner's more substantive rejection as set forth in the Office Action of September 21, 2005 is addressed thereafter in detail, and in the order rejected by the Examiner.

In regard to the summary rejection set forth in the final Office Action of March 24, 2006, the Examiner asserted that in the Amendment of January 4, 2006, the Appellants argued the deficiencies of the cited prior art references individually. However, this is clearly incorrect in that the Appellants described what is actually disclosed in the cited Bowen et al. and England et al. references, which is contrary to what the Examiner asserts they disclose in the Office Action. This clarification as to what the cited prior art discloses was made for the purpose of demonstrating how the Examiner failed to establish a *prima facie* case of obviousness. (See Amendment filed Jan. 4, 2006, Pages 10 to 12). In this regard, the Appellants contended in the Remarks of the amendment that the Examiner has failed to establish a *prima facie* case of obviousness in that 1) there is no motivation in either of these references to combine them in the manner suggested by the Examiner; and 2) even if there was motivation, and the references are combined in the manner suggested, such combination still fail to result in the invention recited in the claims. (See Amendment filed Jan. 4, 2006, Pages 13 to 14. See also MPEP 2142).

Claims 1 and 18

In the prior substantive Office Action mailed September 21, 2005, the Examiner rejected independent claims 1 and 18 (as well as claims 21 and 32 argued separately below) under 35 U.S.C. 103(a) as being unpatentable over Bowen in view of England. Appellants respectfully contend that the Examiner's rejection of independent claims 1 and 18 is improper and should be reversed.

Under 35 U.S.C. 103, a patent may not be obtained if the subject matter sought to be patented would be obvious to a person having ordinary skill in the art to which

the subject matter pertains. A determination of obviousness is a legal conclusion based on underlying findings of fact. *Velander v. Garner*, 348 F.3d 1359, 1363 (Fed. Cir. 2003). The Supreme Court in *Graham v. John Deere*, 383 U.S. 1 at 18, 148 USPQ 459 at 167 (1996), set forth the basic test for patentability under 35 U.S.C. §103:

> Under §103, the scope and content of the prior art are to be determined; differences between the prior art and the claims at issue are to be ascertained; and the level of ordinary skill in the pertinent art resolved. Against this background, the obviousness or non-obviousness of the subject matter is determined. Such secondary considerations as commercial success, long felt but unresolved need, failure of others, etc., might be utilized to give light to the circumstances surrounding the origin of the subject matter to be patented.

Moreover, in *In re Ehrreich* and *Avery*, 200 USPQ 504, 509-510 (CCPA 1979), the Court of Customs and Patent Appeals further clarified the basic test set forth in *Graham v. John Deere*:

> We must not here consider a reference in a vacuum, but against the background of the other references of record which may disprove theories and speculations in the reference or reveal previously undiscovered or unappreciated problems. The question in a §103 case is what the references would <u>collectively suggest</u> to one of ordinary skill in the art. *In re Simon*, 461 F.2d 1387, 174 USPQ 114 (CCPA 1972). It is only by proceeding in this manner that we may fairly determine the scope and content of the prior art according to the mandate of *Graham v. John Deere*, 383 US 1, 17, 148 USPQ 459, 467 (1966)(Emphasis in original).

Thus, "[t]he mere fact that references can be combined or modified does not render the resultant combination obvious unless the prior art also suggests the desirability of the combination," *In re Mills*, 916 F.2d 680, 16 USPQ2d 1430 (Fed. Cir. 1990). Where the prior art provides "only general guidance and is not specific as to the particular form of the invention or how to achieve it, [such a suggestion] may make an approach 'obvious to try,' but it does not make the invention obvious." *Ex parte Obukowicz*, 27 USPQ2d, 1063, 1065 (U.S. Patent and Trademark Office Board of Appeals and Interferences, 1992) and *In re O'Farrell*, 7 USPQ2d 1673, 1681 (Fed. Cir. 1988).

10152339.3

Thus, three criteria must be met to establish a *prima facie* case of obviousness. (see M.P.E.P. §2143.) First, there must be some teaching, suggestion or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988). Second, there must be a reasonable expectation of success. *In re Rhinehart*, 531 F.2d 1048, 189 USPQ 143 (CCPA 1976). Last, the prior art must teach or suggest all the claim limitations. *In re Royka*, 490 F.2d 981, 180 USPQ 580 (CCPA 1974).

In the Office Action mailed September 21, 2005, the Examiner asserted that Bowen teaches a method of extracting data of interest from a plurality of web sites, including creating a description of data of interest (citing Bowen, col. 2, lines 47-66), an extraction pattern identifying at least a portion of the output of a web site and extracting information from the respective web site (citing Bowen, col. 12, lines 5-34), and associating the developed extraction pattern with the respective data of interest (citing Bowen, col. 11, lines 45-67 and col. 12, lines 5-22). The Examiner further asserted that Bowen discloses receiving a value that can be used as an extraction parameter (citing Bowen, col. 12, lines 35-65), and obtaining of the data of interest by querying the web sites, wherein the data includes data of interest from at least two web sites (citing Bowen, col. 11, lines 35-44; col. 2, line 5 to col. 3, line 12; col. 3, line 45 to col. 4, line 3).

In contrast to the assertions above, the Examiner then acknowledges that Bowen does not teach developing an extraction pattern based on output from the respective web site using a graphical user interface tool. (See Office Action mailed September 21, 2005, page 3, last ¶). However, the Examiner notes that Bowen uses a Graphical User Interface (GUI). The Examiner then asserts that England somehow discloses "developing an extraction pattern based on output from the respective web site using a graphical user interface tool." The Examiner cites to Col. 2, line 65 to Col. 3, line 25 of England, and notes various operations a user can perform through the GUI of the browser. The Examiner concludes that it would be obvious to "include graphical user interface (GUI) in a browser to extract information in a pattern format as disclosed in England." (See Office Action mailed September 21, 2005, page 3, last

¶D. The Appellants respectfully disagree with the Examiner's rejection as set forth herein below.

In contrast to the Office Action, the cited portions of the Bowen reference fails to disclose the features asserted by the Examiner. The cited Bowen reference is directed to a method and system for supporting <u>keyword searches</u> of data items in a structured database, such as a relational database, using a GUI. In this regard, the reference discloses that selected data items are retrieved using a SQL query, and the retrieved data values are documented using a markup language such as HTML. (See Bowen, Abstract). The documents are indexed using a web crawler, the indexing agent producing an index that associates keywords with resource locators such as URLs, hot links, file paths, or distinguished names. (See Bowen, Abstract; Col. 8, lines 31-45). When a user provides a keyword query to a search engine interface that incorporates a GUI, the index is used to obtain a resource locator that is associated with the keyword, and used to retrieve the item's current data from the structured database, and to generate the document containing the retrieved data. (See Bowen, Abstract; Col. 8, lines 19-30).

In other words, as clearly taught by the cited portions of Bowen above, as well as those portions relied upon by the Examiner, Bowen merely discloses search engine technology and operation thereof which searches web page indexes for web pages that match the keyword query entered into the search engine. The cited portions of Bowen disclose a database reader that uses SQL queries to extract the exposed data from the database. Of course, various search engine technologies and use of SQL in databases are well known in the art, but this is not the subject of the claimed invention.

The secondary reference to England is directed to a software system which allows a guide/specialist to interact with a client on a real-time, interactive basis over the World Wide Web. (See England, Abstract). England discloses a guide system with a special-purpose browser displaying both locally displayable frames and remotely displayable frames, where the client system utilizes a conventional browser. (See England, Abstract; Col. 7, line 65-Col. 8, line 8). The remotely displayable frames are transmitted to the client so that both the guide and the client have views of the remotely displayable frames. (See England, Abstract; Col. 8, lines 53-65).

Correspondingly, the England reference discloses a customized browser which allows interaction with a client system. Of course, various browser technologies are well known in the art, but is also not the subject of the claimed invention.

In the above regard, the Examiner fails to establish proper motivation for combining Bowen and England together in the Office Action, and thus, does not establish a *prima facie* case of obviousness. In particular, there is no motivation established in either Bowen nor England references to combine them, and to modify the system of Bowen using the teachings of the cited England reference, as suggested by the Examiner. This is expected since these cited references are directed to two different technological fields that are not related, Bowen being directed to system for supporting keyword searches of data items in a database such as a relational database, while England is directed to a guide system with a special-purpose browser displaying both locally displayable frames and remotely displayable frames.

The Examiner fails to identify any teachings in either references to combine them in any manner, but merely asserts that "it would have been obvious to one of ordinary skill in the art to include graphical user interface (GUI) in a browser to extract information in a patterned format" without any basis. (See Pg. 4, first ¶). In addition, extracting information in a "patterned format" is not relevant to the present invention which is directed to extracting information using an extraction pattern that is developed using the output of a web site. Because there is no teachings or suggestions in either of the cited references to combine the references in the manner suggested by the Examiner, the Appellants respectfully contend that the Examiner has failed to properly establish a *prima facie* case of obviousness. Correspondingly, the reversal of the present rejection is respectfully requested.

Secondly, even if there was motivation to combine these references in the manner suggested by the Examiner, they still fail to result in the present invention recited invention of claims 1 and 18. In particular, even if the system disclosed in Bowen was modified to have a customized graphical user interface as disclosed in the cited England reference, such combination would still fail to result in a system as claimed that allows developing of an extraction pattern based on output from the respective web pages as noted above. In particular, independent claim 1 specifically

recites developing an extraction pattern from a <u>web page output</u> from the respective website using a graphical user interface tool, and that the extraction pattern is adapted to identify at least a portion of an output a web site and <u>to extract information from a plurality of web pages</u> of the respective web site. In other words, as previously described, rather than creating programs such as crawlers that extract information from web pages of a web site as taught in the prior art, the present invention allows the output of the web site to be used to generate extraction patterns that are then used to extract the desired information from the web pages of the web site.

Mere search engines such as that described in Bowen cannot perform the function of the present invention, and do not render the present invention as claimed unpatentable. For instance, in the system of Bowen, the particulars for each book, such as a title, author, or other parameter would be required in order to retrieve information about books from a web site. As previously noted, extraction patterns of the present invention and recited in the present claims are not the same, or even equivalent to, keyword queries. Thus, even if the system disclosed in Bowen was modified to have a customized graphical user interface as disclosed in the cited England reference, such combination would still fail to result in a system as claimed that allows developing of <u>an extraction pattern based on the output</u> from the respective web site, <u>or extracting of information from a plurality of web pages from a web site using the developed extraction pattern</u>. Moreover, as also noted above, extracting information in a patterned format is not claimed. Instead, development of extraction patterns using an output of the website is claimed. Correspondingly, the reversal of the Examiner's rejection of claims 1 and 18 based on Bowen and England is also requested for the above reason as well.

<u>Claim 21</u>

Regarding Examiner's rejection of independent claim 21, the comments set forth above relative to independent claims 1 and 18 are equally applicable. Moreover, claim 21 further recites operations for testing the instruction using the extraction pattern and the contents of a buffer (Fig.4; 414; Pg. 19, lines 3-6), and that the buffer includes a portion of the web page output of the web site associated with the

description of data of interest (Pg. 16, lines 7-8; Fig. 10, 1002; Pg. 23, lines 13-19). These limitations in claim 21 are not disclosed in the cited prior art and not addressed in the Examiner's rejection. Correspondingly, the Examiner's rejection of independent claim 21 is believed to be improper, and the Appellants request the reversal thereof.

### Claim 32

Regarding Examiner's rejection of independent claim 32, the comments set forth above relative to independent claim 1 is equally applicable. Moreover, claim 32 further recites that the graphical user interface tool includes a web browser (Pg. 14, lines 13-15; Fig. 9; Pg. 22, line 22-Pg. 23, line 2). This limitation in claim 32 is not addressed in the Examiner's rejection. Correspondingly, the Examiner's rejection of independent claim 32 is believed to be improper, and the Appellants request the reversal thereof.

### Claims 2, 7, and 20

In the Office Action, dependent claims 2, 7, and 20 were rejected based on England, the Examiner simply noted that England discloses a GUI tool that includes a web browser. However, this rejection is believed to be improper in view of their dependency on independent claim 1 (claims 2 and 7) or independent claim 18 (dependent claim 20), and the reversal of the Examiner's rejection is respectfully requested.

### Claim 5

Dependent claim 5 was also rejected in the Office Action, the Examiner simply noting that England discloses a GUI tool that includes a web browser. However, claim 5 is believed to be allowable at least for the reason of its dependency on dependent claim 2, and ultimately on claim 1. Moreover, neither England nor Bowen disclose or suggest a plurality of predefined extraction patterns as recited in claim 5, since extraction patterns are not utilized in the systems of England and Bowen. No further basis for this rejection is set forth by the Examiner in the Office

Action. Correspondingly, this rejection is believed to be improper and the reversal of this rejection is also respectfully requested.

Claim 4

Dependent claim 4 was rejected based on Bowen. However, this claim is dependent on claim 2, and ultimately dependent on independent claim 1. Thus, claim 4 is believed to allowable at least for the reason of its dependency.

Moreover, dependent claim specifically recites applying the extraction pattern to the output of the web site that is displayed in a source view in the web browser thereby identifying the at least a portion of the output for the web site (Pg. 16, lines 7-8; Fig. 10, 1002; Pg. 23, lines 13-19), and displaying a rendered version of the at least a portion of the output of the web site (Pg. 16, lines 8-9; Fig. 14; Pg. 24, lines 18-22).

In contrast to the Examiner's assertion, Bowen does not disclose or suggest application of an extraction pattern to the output of the web site displayed in a source view. The cited portion of the Bowen reference (Col. 13, lines 5-52) relied upon by the Examiner merely describes operation of indexing technology that utilizes SQL to search for, and retrieve, documents that match the query. The reference fails to disclose application of an extraction pattern to the output of a web site, much less an output displayed in a source view. Furthermore, just because different views may be possible in prior art systems does not mean that an extraction pattern is applied to a particular view as recited. Clearly, this summary rejection of claim 4 by the Examiner is improper and the Appellants request the reversal thereof.

Claim 9

Dependent claim 9 is dependent on claim 1, and thus, is believed to be allowable at least for the basis of its dependency. Moreover, dependent claim 9 recites that the developing of an extraction pattern includes receiving a selection of an extraction command from a predetermined list of extraction commands (Fig. 4, 406; Pg. 16, lines 13-16; Pg. 17, Table 1). The cited Bowen reference fails to disclose or suggest receiving a selection of an extraction command from a predetermined list of extraction commands. Again, the cited portion of the Bowen reference (Col. 13, lines

5-52) relied upon by the Examiner merely describes operation of indexing technology that utilizes SQL to search for, and retrieve, documents that match the query, and does not appear to be relevant to the claimed invention. Thus, this rejection is believed to be improper and the Examiner's rejection of this claim should be reversed as well.

### Claim 12

Dependent claim 12 is dependent on dependent claim 9 discussed above, and thus, is believed to be allowable at least for the reason of its dependency. Moreover, dependent claim 12 recites applying a test condition comprising a logical test for at least one corresponding argument (Fig.4; 414; Pg. 19, lines 3-6). The cited Bowen reference fails to disclose or suggest this limitation, and no additional basis in support of this rejection is provided by the Examiner. Thus, this rejection is believed to be improper, and rejection is respectfully requested.

### Claim 14

Dependent claim 14 is dependent on claim 12 discussed above, and thus, is believed to be allowable at least for the reason of its dependency. Moreover, dependent claim14 recites that the test condition further comprises a result code that returns an error when the output of the respective web site has changed (Pg. 20, lines 20-25). The cited Bowen reference fails to disclose or suggest a feature or function, and no additional basis in support of this rejection is provided by the Examiner. Thus, this rejection is believed to be improper and the Examiner's rejection of this claim should also be reversed.

### Claim 15

Dependent claim 15 is also dependent on claim 12 discussed above, and thus, is believed to be allowable at least for the reason of its dependency. Moreover, dependent claim 15 recites that the test condition further comprises a result code that returns an error when the output of the respective web site has no information about the product (Pg. 20, lines 14-19). The cited Bowen reference fails to disclose or suggest a feature or function, and no additional basis in support of this rejection is provided by the Examiner. Thus, this rejection is believed to be improper and the

reversal of the Examiner's rejection is respectfully requested.

### Claim 16

Dependent claim 16 is dependent on claim 9 discussed above, and thus, is believed to be allowable at least by the reason of its dependency. Moreover, dependent claim 16 recites that the predetermined list of extraction commands includes an extraction command for segmenting the output of the respective web site into a plurality of units, each of the plurality of units matching the extraction pattern (Pg. 19, lines 20-24). The cited Bowen reference fails to disclose or suggest a feature or function, and no additional basis in support of this rejection is provided by the Examiner. Thus, this rejection is believed to be improper, and the Examiner's rejection of this claim should also be reversed.

### Claim 6

Dependent claim 6 is dependent on claim 5 discussed above, and thus, is believed to be allowable at least for the reason of its dependency. In addition, claim 6 recites that the plurality of predefined extraction patterns includes at least one of an extraction pattern for matching a hyperlink, an extraction pattern for matching a form, and an extraction pattern for matching a price (Fig. 10, 908). In the Office Action, this claim was also rejected based on Bowen. However, the basis for this rejection is wholly unclear in that the cited portion of the Bowen reference merely discloses various types of documents that may include links, for example. Of course, such documents are known, but the Bowen reference does not disclose predefined extraction patterns of the present invention that perform particular function as recited in claim 6. Thus, this rejection is believed to be improper and the reversal of the Examiner's rejection is respectfully requested.

### Claims 25 and 26

Dependent claims 25 and 26 are dependent on independent claim 1 discussed above. Thus, these claims are believed to be allowable at least by the reason of their dependency, and the reversal of the Examiner's rejection is requested.

Claim 19

Claim 19 is dependent on independent claim 18 discussed above, and thus, is believed to be allowable at least for the reason of its dependency. Claim 19 also recites that the means for developing an extraction pattern includes means for selecting an instruction from a predetermined list of instructions (Fig. 4, 406; Pg. 16, lines 13-16; Pg. 17, Table 1). Claim 19 was rejected by the Examiner based on Bowen, but the basis for this rejection is not understood since the cited portions of the Bowen reference does not disclose an extraction pattern or a predetermined list of commands at all, but instead discusses hardware of a computer and various programmable languages. Thus, the Appellants contend that this rejection is improper and should be reversed.

Claims 23, 24, 27, 28 and 33

Dependent claims 23, 24, 27, 28 and 33 were also rejected by the Examiner citing Bowen. Claims 23 and 24 are ultimately dependent on independent claim 1, claim 27 is dependent on independent claim 18, claim 28 is dependent on independent claim 21, and claim 33 is dependent on independent claim 32. Thus, these claims are believed to be allowable at least for the reason of their dependency on allowable independent claims.

In addition, these claims recite that the extraction pattern comprises a pre-condition regular expression, a portion of data of interest regular expression, and a post-condition regular expression, and the refining the pre-condition regular expression, the portion of data of interest regular expression, or the post-condition regular expression (Pg. 15, lines 13-24).

The portions of Bowen cited by the Examiner only relates to keyword searching and use of wildcards in such keyword searches. The cited portions do not relate at all to the recited extraction patterns that are developed from a web page output of the web site, or the recited refining of the extraction pattern. Thus, this rejection is improper and the reversal of this rejection of claims 23, 24, 27, 28 and 33 is respectfully requested.

Claims 29-31 and 34

Dependent claims 29-31 and 34 were also rejected by the Examiner based on Bowen. However, claim 29 is dependent on independent claim 1, claim 30 is dependent on independent claim 18, claim 31 is dependent on independent claim 21, and claim 34 is dependent on independent claim 32. Correspondingly, these claims are believed to be patentable at least for the reason of their dependency. Therefore, the reversal of this rejection of these claims is requested.

Claims 10, 11, 13 and 17

Dependent claims 10, 11, 13 and were not specifically addressed by the Examiner in the Office Action. Correspondingly, no arguments can be presented except to note that these claims are ultimately dependent on dependent claim 9 discussed above, and ultimately depend on independent claim 1 also discussed above. Therefore, these claims are believed to be allowable at least for the reason of their dependency, and the reversal of the Examiner's rejection is requested.

## VIII.  CLAIMS APPENDIX

1. A method of extracting data of interest from a plurality of web sites, the method comprising:

(A) for each respective web site W in said plurality of web sites,

(i) creating a respective description of data of interest that identifies the web site W;

(ii) developing an extraction pattern from a web page output from the respective web site using a graphical user interface tool, the extraction pattern being adapted to identify at least a portion of an output of a web site and to extract information from a plurality of web pages of the respective web site W; and

(iii) associating the developed extraction pattern with the respective description of data of interest for the respective web site W;

(B) receiving a value for use as an extraction parameter for the developed extraction patterns; and

(C) obtaining said data of interest by querying web sites in the plurality of web sites using the value and the extraction patterns associated with the respective descriptions of data of interest, wherein

when the data of interest includes data of interest from at least two web sites of the plurality of web sites, the data of interest from the at least two web sites is extracted.

2. The method of claim 1, wherein the graphical user interface tool includes a web browser.

4. The method of claim 2, further comprising:

applying the extraction pattern to the output of the web site that is displayed in a source view in the web browser thereby identifying the at least a portion of the output for the web site; and

displaying a rendered version of the at least a portion of the output of the web site.

5. The method of claim 2, wherein the graphical user interface tool further includes a plurality of predefined extraction patterns.

6. The method of claim 5, wherein the plurality of predefined extraction patterns includes at least one of an extraction pattern for matching a hyperlink, an extraction pattern for matching a form, and an extraction pattern for matching a price.

7. The method of claim 2, wherein the graphical user interface tool further allows:

identifying a form in the output of the respective web site;

submitting the form while creating the description of data of interest corresponding to the web site without retrieving the web site;

generating a plurality of parameters corresponding to an input in the form; and

associating a parameter in the plurality of parameters with the extraction parameter.

9. The method of claim 1, wherein the developing of an extraction pattern includes receiving a selection of an extraction command from a predetermined list of extraction commands.

10. The method of claim 9, wherein the predetermined list of extraction commands includes an extraction command for retrieving multiple matches of an extraction pattern from a web site.

11. The method of claim 9, wherein the predetermined list of extraction commands includes an extraction command for extracting data from a first web site and a second web site, the first web site including a reference to the second web site.

12. The method of claim 9, further including applying a test condition comprising a logical test for at least one corresponding argument, and wherein the respective description of data of interest continues executing when the logical test is satisfied.

13. The method of claim 12, wherein the at least one corresponding argument includes an extraction pattern.

14. (The method of claim 12, wherein the test condition further comprises a result code that returns an error when the output of the respective web site has changed.

15. The method of claim 12, wherein the test condition further comprises a result code that returns an error when the output of the respective web site has no information about the product.

16. The method of claim 9, wherein the predetermined list of extraction commands includes an extraction command for segmenting the output of the respective web site into a plurality of units, each of the plurality of units matching the extraction pattern.

17. The method of claim 16, wherein developing an extraction pattern includes using an extraction command to segment the web page output of the respective web site into a plurality of units, and using a test condition that comprises a logical test and at least one argument, and wherein for each of the plurality of units, the logical test is computed with the at least one argument, and the unit is removed from the plurality of units if the logical test is not satisfied with the at least one argument.

18. An apparatus for extracting information of interest from a plurality of web sites, the apparatus comprising:

(A) for each respective web site W in the plurality of web sites,

(i) means for creating a respective description of data of interest that identifies the web site W;

(ii) means for developing an extraction pattern from a web page output from the respective web site using a graphical user interface tool, the extraction

pattern being adapted to extract data from a plurality of web pages of the web site; and

(iii) means for associating the developed extraction pattern with the respective description of data of interest for the respective web site W;

(B) means for receiving a value for use as an extraction parameter in the developed extraction patterns; and

(C) means for obtaining said data of interest by querying web sites in the plurality of web sites using the value and the developed extraction patterns associated with the respective descriptions of data of interest,

wherein, when the data of interest includes data from at least two web sites of the plurality of web sites, the means for obtaining said data of interest provides the data of interest from the at least two web sites.

19. The apparatus of claim 18, wherein the means for developing an extraction pattern includes means for selecting an instruction from a predetermined list of instructions.

20. The apparatus of claim 18, wherein the graphical user interface tool comprises a web browser.

21. A computer data signal embodied in a carrier wave comprising:

(A) a software module for creating a description of data of interest, the software module including;

(i) a set of operations for interactively developing an extraction pattern from a web page output of a target web site using a graphical user interface tool, the developed extraction pattern being adapted to extract data of interest from a plurality of web pages of the target web site;

(ii) a set of operations for receiving a selection of an instruction from a predefined set of instructions for inclusion in the description of data of interest;

(iii) a set of operations for associating the extraction pattern with the instruction;

(iv) a set of operations for testing the instruction using the extraction pattern and the contents of a buffer, wherein the buffer includes a portion of the web page output of the web site associated with the description of data of interest; and

(B) a software module for using said description of data of interest to obtain data of interest from the target web site when a value for use as an extraction parameter for the developed extraction pattern is provided.

23. The method of claim 1 wherein said extraction pattern comprises a pre-condition regular expression, a portion of data of interest regular expression, and a post-condition regular expression and wherein said operations for interactively developing an extraction pattern comprises refining at least one of said pre-condition regular expression, said portion of data of interest regular expression, and said post-condition regular expression.

24. The method of claim 23 wherein said portion of data of interest regular expression includes a variable that is replaced with said value for said extraction parameter.

25. The method of claim 1 wherein the data of interest is provided incrementally as it is obtained from the plurality of web sites.

26. The method of claim 1 wherein, the data of interest is obtained from the plurality of web sites and then presented simultaneously.

27. The apparatus of claim 18 wherein said extraction pattern comprises a pre-condition regular expression, a portion of data of interest regular expression, and a post-condition regular expression and wherein said means for developing comprise refining at least one of said pre-condition regular expression, said portion of data of interest regular expression, and said post-condition regular expression.

28. The computer data signal of claim 21 wherein said extraction pattern comprises a pre-condition regular expression, a portion of data of interest regular expression, and a post-condition regular expression and wherein said operations for developing comprise refining at least one of said pre-condition regular expression, said portion of data of interest regular expression, and said post-condition regular expression.

29. The method of claim 1 wherein the data of interest is information associated with a product or information associated with a service.

30. The apparatus of claim 18 wherein the data of interest is information associated with a product or information associated with a service.

31. The computer data signal of claim 21 wherein said data of interest is a product, information, or a service.

32. A computer implemented method of obtaining data of interest from a plurality of web sites comprising:

(A) developing a description of data of interest for each web site in said plurality of web sites from web page output from the plurality of web sites using a graphical user interface tool that includes a web browser, each respective description of data of interest specifying an address for a corresponding web site in the plurality of web sites and each respective description of data of interest including an extraction pattern adapted to identify at least a portion of the output of a web site and to extract user specified information from a plurality of web pages of the corresponding web site;

(B) receiving a value for use as an extraction parameter for the developed extraction patterns; and

(C) obtaining said data of interest by querying web sites in the plurality of web sites using the value and the extraction patterns in the respective descriptions of data of interest.

33. The computer implemented method of claim 32 wherein each said extraction pattern comprises a pre-condition regular expression, a portion of data of

interest regular expression, and a post-condition regular expression and wherein said step of developing a description of data of interest comprises refining at least one of said pre-condition regular expression, said portion of data of interest regular expression, and said post-condition regular expression.

34. The computer implemented method of claim 32 wherein said data of interest is a product, information, or a service.

## IX.    EVIDENCE APPENDIX
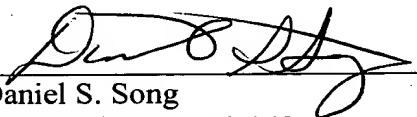
None.

## X.  <u>RELATED PROCEEDINGS APPENDIX</u>

None.

## XI.   CONCLUSION

Thus, at least for the foregoing reasons, the Appellants contend that the Examiner's rejection of the presently pending claims is improper in that the cited Bowers and England references, in combination, do not render the claimed invention obvious or unpatentable.  Therefore, the reversal of the Examiner's rejection under 35 U.S.C. §103(a) with respect to all of the pending claims 1, 2, 4-7, 9-21 and 23-34 are respectfully requested.

Respectfully submitted,

Date:   March 28, 2007

Daniel S. Song
Registration No. 43,143

**Customer No.:  22204**

**NIXON PEABODY LLP**
401 9th Street, N.W., Suite 900
Washington, DC  20004
(202) 585-5000
(202) 585-8080 (Fax)